medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Demarcation line determination for diagnosis of gastric cancer disease range using unsupervised machine learning in magnifying narrow-band imaging

Shunsuke Okumura[1], Satoru Hiwa[2]*, Takeshi Yasuda[3], Hiroaki Kitae[3], Yuriko Yasuda[3], Akira Tomie[3], Tatsushi Omatsu[3], Hiroshi Ichikawa[4], Nobuaki Yagi[3], and Tomoyuki Hiroyasu[2]

[1]Graduate School of Life and Medical Sciences, Doshisha University, Kyoto, Japan
[2]Department of Biomedical Sciences and Informatics, Doshisha University, Kyoto, Japan
[3]Department of Gastroenterology, Asahi University Hospital, Gifu, Japan
[4]Department of Medical Life Systems, Doshisha University, Kyoto, Japan

***Correspondence**: Satoru Hiwa, Department of Biomedical Sciences and Informatics, Doshisha University, 1-3 Tatara Miyakodani, Kyotanabe, Kyoto 610-0394, Japan; shiwa@mail.doshisha.ac.jp

Disclosures:
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abstract:
Background and aims: It is important to determine an accurate demarcation line (DL) between the cancerous lesions and background mucosa in magnifying narrow-band imaging (M-NBI)-based diagnosis. However, it is difficult for novice endoscopists. Recently, machine learning (ML) approaches such as deep learning have been applied to this issue; however, this results in the burden of training ML models and labeling large datasets. To overcome these issues, we propose and a DL determination system using an unsupervised ML method.
Methods: Our method consists of the following four steps: 1) An M-NBI image is segmented into superpixels (a group of neighboring pixels) using simple linear iterative clustering. 2) The image features are extracted for each superpixel. 3) The superpixels are grouped into several clusters using k-means method. 4) The boundaries of the clusters are extracted as DL candidates. To validate the proposed method, 16 M-NBI images of 4 cases were used for performance evaluation. The evaluation investigated the similarity of the DLs identified by endoscopists and our method, and the Euclidean distance between the two DLs was calculated.
Results: The average Euclidean distances for the four cases were 13.07, 19.95, 12.53, and 17.90. The results indicated that the specific selection of the number of clusters enabled the proposed method to detect DLs that were similar to those of the endoscopists. The DLs identified by our method represented the complex shapes of the DLs, similarly to those identified by experienced doctors.
Conclusion: Our proposed system can support the training of inexperienced doctors, as well as enriching the knowledge of experienced doctors in endoscopy.

Keywords: computer-aided diagnosis, unsupervised machine learning, demarcation line, magnifying narrow-band imaging, early gastric cancer

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

1. INTRODUCTION

Gastric cancer is the third leading cause of cancer-related deaths following lung cancer and colorectal cancer, and has claimed the lives of approximately 729,000 people worldwide per year [1]. The five-year survival rate exceeds 90% if gastric cancer is detected at an early stage, but it declines in later cancer stages [2], [3]. Thus, early detection is a crucial issue in gastric cancer treatment.

Several studies have reported that magnifying narrow-band imaging (M-NBI) is useful for diagnosing the existence of early gastric cancer and determining its horizontal extent [4], [5], [6]. The detected cancer lesions can subsequently be removed endoscopically by means of endoscopic submucosal dissection (ESD) [7], [8].

However, ESD exhibits certain matters of concern, such as complications and the risk of non-curative resection (remnants of cancer) [9], [10], [11]. Although this is also the case for other endoscopic procedures, accidental perforation sometimes occurs [12], [13], [14]. Furthermore, the incorrect recognition of the demarcation line (DL) is disadvantageous to patients. When endoscopists underestimate the DL, this may lead to non-curative resection. In contrast, if they overestimate the DL, they may perform ESD containing extensive unnecessary lesions, which also increases the risk of perforation or a long procedure time. Therefore, it is important to determine an accurate DL between the cancerous lesions and background mucosa. However, this determination is difficult for novice endoscopists, with long-term training and experience being required.

In this context, with the recent development of machine learning (ML) technology, computer-aided diagnosis (CAD) systems are playing an increasingly important role in the medical field, and are expected to aid endoscopists with accurate and stable disease detection and diagnosis. In the gastroenterology field, CAD systems for the detection of esophagus cancer, gastric cancer, *H. pylori* infection, and the characterization of colorectal lesions have been developed and validated for clinical use [15], [16], [17], [18].

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In this study, we developed a new CAD system that automatically determines the DL on a given M-NBI image through an ML method.
The main objectives of the proposed system are to provide the DL candidates for diagnosis of the disease range of early gastric cancer to endoscopists, as well as to support accurate and stable diagnosis. In recent years, research on image diagnosis using supervised ML such as deep learning has attracted significant interest in medicine owing to its higher classification and prediction performance compared to conventional methods [15], [16], [19], [20]. However, disadvantages exist in obtaining high-performance results in these methods. First, it is necessary to prepare an enormous amount of annotated training data (for example, 'lesion' or 'non-lesion' labels are provided to each datum by endoscopists) to construct the classification model. These annotations need to be performed by highly experienced endoscopists, and are therefore impractical. Second, the diagnosis rationale and criteria are unclear because the available deep learning methodologies are mainly black-box models. It is difficult to extract such knowledge from a deep learning model [17].

To overcome these issues, in this study, we used a data clustering method, which is a type of unsupervised ML technique. Given a single M-NBI image, the proposed method divides it into several segments, and then identifies the clusters in which segments with relevant features are densely located. Consequently, the boundary between the gastric cancer lesions and background mucosa; that is, the DL, can be identified. The diagnosis rationale or criteria can be interpreted following the determination of the DL because we specify the potentially discriminating features of gastric cancer prior to performing the data clustering. Furthermore, the proposed method does not require any annotations or computational efforts. Thus, the approach can be applied to determine the DL for early gastric cancer diagnosis.

2. METHODS

2.1 Proposed method

A schematic of the proposed method is presented in **Figure 1**. It mainly consists

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

of the following four steps: (1) image segmentation, (2) feature extraction, (3) segment-based clustering, and (4) identification of DL candidates. First, an input M-NBI image is segmented into a certain number of superpixels, each of which is a set of pixels with similar color features. Second, the image features including not only the colors but also the texture features are calculated for each superpixel. Third, the superpixels are grouped into a predetermined number of clusters using a data clustering method. Finally, the boundaries of the multiple clusters are extracted as the DL candidates and are presented to the users of the proposed system (that is, endoscopists). The majority of data clustering methods divide the dataset into clusters so as to minimize the intracluster distance or/and maximize the intercluster distance. Therefore, we hypothesized that the DL could be extracted as the boundary of the clusters, because the cancer lesions that would have different image features from the non-cancer lesions should form the cluster. Note that our system was designed to identify only the DL 'candidates', which is why it does not require image annotation by endoscopists. The final decision of selecting the DL from the candidates is the responsibility of the user endoscopists. The details of each of the four steps are described in the following sections.

2.1.1 Image segmentation

Several superpixel algorithms have been presented that partition an image into multiple segments. In this study, we used simple linear iterative clustering (SLIC) [21], which is one of the most popular methods in image processing. In the SLIC algorithm, a group of pixels that are similar in terms of both their color features and position (coordinates) are formed as a superpixel. The boundary shapes of the superpixels are iteratively refined by means of the well-known *k*-means-based clustering approach. First, the input image is divided into a certain number of grids. The center of each tile (the initial form of a superpixel) is sampled and then moved to the pixel with the lowest color gradient among the neighboring pixels. Thereafter, for each pixel in the entire image, the nearest center pixel in terms of both the color and coordinate spaces is searched, and it is assigned to the member of the identified pixel. This procedure is repeated until certain types of termination criteria are satisfied, so that the superpixels are

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

generated. For further details on SLIC, please refer to the original paper [21]. In this study, we used the *slic* function implemented in scikit-image, which is a Python library.

Prior to applying the SLIC image segmentation, we had to determine the color space to represent the image feature. In this study, we employed the International Commission on Illumination (CIE) L*a*b* color space as the image features for segmentation because it is device independent and also matches with human color perception. It is represented by three-dimensional space consisting of L*, which represents the darkness to lightness, a*, which varies from red to green, and b*, which varies from blue to green [22].

2.1.2 Feature extraction and clustering

After segmenting an input M-NBI image into a set of superpixels, they are further merged into the clusters of the superpixels using the *k*-means method. That is, the local clustering of the pixels to extract the pixel-level image feature is performed by SLIC, following which the global clustering of the superpixels is further performed by the *k*-means method. In the global clustering process, the aim is to identify a cluster of cancer lesions by evaluating each superpixel in terms of both the color and texture features.

The CIE L*a*b* and HSV color spaces were used for representing the color features. The HSV color space is defined by three components: hue, saturation, and value. Hue represents the color type (for example, red, blue, and yellow) ranging from 0 to 360 degrees, saturation represents the color vividness ranging from 0 to 100%, and value represents the color brightness from 0 to 100% [23]. The coordinates in both color spaces can be calculated by nonlinear transformation from the RGB color space.

In addition to the color features, we aimed to extract the texture features of gastric cancer lesions. Yao et al. [24] set their criterion for early gastric carcinoma as (1) the presence of an irregular microvascular pattern or (2) the presence of an irregular microsurface pattern, both of which were accompanied

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

by a clear DL. Therefore, we expected that the texture features of a gastric cancer lesion could be extracted if the irregularities of the microvascular and microsurface patterns were quantified. We used the local entropy, which quantifies the complexity such as the randomness of the image, proposed by Wu et al. [25]. The local entropy score *H*(*S*) of the image block *S* (corresponding to a superpixel in this study) in a given image can be defined by:

$$H(S) = -\sum_{l=0}^{L-1} P_l \log_2 P_l = -\sum_{l=0}^{L-1} \frac{n_l}{T} \log_2 \frac{n_l}{T}, \tag{1}$$

where $l \in \{0, 1, \ldots, L-1\}$ denotes the pixel intensity scale (for example, $L = 256$ for an 8-bit grayscale image) and n_l, T, and P_l denote the total number of pixels within S at pixel intensity l, the total number of pixels, and the probability of a pixel having the intensity l, respectively. In brief, the entropy used in this case quantifies the variability of the pixel values in the image. That is, we assumed that the irregularities of the microvascular and microsurface patterns could be represented by the probability distribution of the pixel values. Note that we calculated the local entropy score for each color dimension.

In total, we used 12-dimensional feature vectors (that is, the six color dimensions L*, a*, b*, H, S, and V and the local entropy score for each of the six color dimensions) for the global clustering by the *k*-means method. We used the *KMeans* function implemented in scikit-learn.

2.1.3 Identification of DL

Following the global clustering of the superpixels, the boundaries of the clusters are extracted as the DL candidates, and are presented on the graphical user interface for the user endoscopists. As mentioned previously, our system is designed to provide the DL candidates and not to determine the exact DL. The final decision on the DL determination is made by the endoscopists by selecting one of the candidates. In this situation, it is necessary to provide varieties of the candidates for the users. Therefore, global clustering with different numbers of clusters (the control parameter *k* of the *k*-means method) was executed to provide the DL candidates at different segmentation levels of the M-NBI image.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Examples of the cluster boundaries identified following the global clustering is illustrated in **Figure 2**.

2.2 Experimental setup

To evaluate how accurately our system could provide the DL candidates, we compared them with DLs identified by endoscopists.

2.2.1 Test images

Endoscopic images obtained from October 2014 to January 2015 was used to establish the artificial intelligence (AI) system for the diagnosis of the disease range of gastric cancer. Those images were anonymized for the sake of personal information is not specified. This study was approved by the ethics committee of the Asahi University Hospital (registration number: 2020-05-01) and was conducted in accordance with the Helsinki Declaration of the World Medical Association. In this study, advanced expertise was required for processing images and making diagnostic algorithms. Therefore, all patients or the public were not involved in the design, conduct, reporting, and dissemination plans of our research.

A total of 16 gastric M-NBI images of 4 cases were used for the performance evaluation. All examinations were carried out using the EVIS LUCERA or EVIS LUCERA ELITE system and the upper gastrointestinal endoscope GIF-FQ260Z (Olympus Medical Systems, Co., Ltd., Tokyo, Japan). These endoscopic systems have normal observation and NBI modes. All images were acquired in the NBI mode combined with magnifying endoscopy. The clinical characteristics of the patients with early gastric cancer are summarized in **Table 1**. The resolutions of these images were of two types: 865 × 999 and 736 × 850. Moreover, board-certified gastroenterologists at Asahi University Hospital had drawn the DLs on these images. An example of a DL drawn by an endoscopist is presented in **Figure 3**.

2.2.2 Parameter settings of proposed method

For the SLIC during the local clustering process, the number of segments

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

(superpixels) of a given image, '*compactness*' parameter, which controls the balance between the color and space proximity, and '*sigma*' parameter, which indicates the width of the Gaussian smoothing kernel for each color dimension, were set to 100, 5, and 5 for L*, a* and b*, respectively. Moreover, for the global clustering process, the maximum number of iterations of the *k*-means algorithm for a single run was set to 300. The initial cluster centers were determined by the *k-means*++ algorithm to accelerate convergence. The execution was carried out with different *k* settings: *k* = 2, 3, 4, 5, and 6.

2.2.3 Performance metrics

We aimed to investigate the similarity between the DLs identified by endoscopists and the proposed method. Therefore, we calculated the Euclidean distance from each pixel of the DL provided by the endoscopist to the nearest pixel of the DL candidates identified by the proposed method as a performance index. **Figure 4** depicts the Euclidean distance calculation between the DLs (DL candidates) provided by the endoscopists and the proposed method. As we used two different image sizes, all of the image sizes were adjusted to the smaller one (736 × 850) to enable a fair comparison in the Euclidean distance calculation.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 1. Clinical features of patients included in test images

Characteristic	Value
Male/female, *n*	3/1
Age, mean, (range), years	71 (59–76)
Tumor size, mm	30
	8
	14
	16
Tumor location, *n*	
Upper	0
Middle	3
Lower	1
Morphologic types, *n*	
Type 0–I	0
Type 0–IIa	1
Type 0–IIb	2
Type 0–IIc	1
Depth of tumor, *n*	
T1a (mucosa)	3
T1b (submucosa)	1
T2 (muscularis propria)	0
Histological classification, *n*	
Differentiated type	4
Undifferentiated type	0

3. RESULTS

Figure 5 presents the DL candidates provided by the proposed method with different settings of the number of clusters (k = 2, 3, 4, 5 and 6) for each of the 16 M-NBI images obtained from 4 cases. The leftmost image includes the DL provided by the endoscopist. The Euclidean distance between the true DL and

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

DL candidates provided by the proposed method is indicated below each image at different *k* values.

Moreover, the images including the nearest DL candidate to the DL drawn by the endoscopist (that is, where the proposed method provided the best performance) are enclosed by blue lines, whereas the maximum ones are enclosed by red lines. Notably, the average minimum Euclidean distance for CASE 1, CASE 2, CASE 3, and CASE 4 were 13.07, 19.95, 12.53, and 17.90, respectively.

Figure 6 indicates the dependence of the Euclidean distance between the DLs provided by the endoscopists and the proposed method on the number of clusters. The dashed black lines correspond to all of the images for all cases. The average distance for all images at each *k* value was calculated and is plotted with a blue point, and these points are connected with a dashed blue line. The average values for $k = 2, 3, 4, 5$, and 6 were 274.47, 113.54, 22.95, 19.20, and 17.80, respectively.

4. DISCUSSION

It can be observed from **Figure 5** that the different cluster number settings resulted in varying shapes of the DL candidates, and the distance to the DL provided by the endoscopist varied accordingly. Furthermore, **Figure 6** indicates that the DL candidates provided by the proposed method became more similar to the endoscopist identification with an increase in the number of clusters. In most cases, the changes in the Euclidean distance were remarkable when *k* ranged from 2 to 4, and were saturated at $k >= 4$. The approximation accuracy was improved with an increase in the number of clusters, because the complexity of the boundary shape increased. These results suggest that the proposed method could identify DL candidates that sufficiently approximated the endoscopist criteria at $k = 4$.

To confirm the above assumption, we observed several typical cases of the identified DL candidates. **Figure 7** presents the DL candidates identified by the

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

proposed method for CASE 4. At $k = 2$, the cluster boundary was drawn between the field of view (FOV) and black corner. At $k = 3$, although the boundary was added in the FOV, it simply divided the FOV into two regions where the brightness of the clusters were differentiated. At $k = 4$, the boundary was finally drawn between the clusters with different mucosal patterns. It can be observed that the DL candidate at $k >= 4$ overlapped the DL provided by the endoscopist. These results visually confirm the assumption that our proposed method could sufficiently approximate the endoscopist criteria at $k = 4$.

However, as demonstrated by CASE 3 in **Figure 8**, when no shaded regions were observed in the M-NBI image, the proposed method approximated the DL drawn by the endoscopist sufficiently at $k = 3$. No further improvements were observed at $k > 3$, as confirmed both visually and quantitatively in **Figure 8**.

Therefore, it is confirmed that our proposed method could identify the DL similar to that provided by the endoscopist at approximately the number of clusters $k = 4$. Obviously, as this process depends on the varieties of M-NBI images, user endoscopists would select the appropriate one by adjusting the k value themselves. However, the observation of multiple DL candidates at different k values may offer new insights into endoscopic diagnosis to user endoscopists, because our method can indicate DLs in regions where endoscopists cannot provide a clear demarcation. This is one of the most important contributions of our study.

A further contribution of our study is that it provides the boundary 'line' itself, instead of indicating the cancer lesions as a 'region'. For example, Li et al. [26] divided M-NBI images into non-cancerous and early gastric cancer groups using a convolutional neural network, which is one of the most well-known deep learning models. Moreover, Kanesaka et al. [27] detected cancer lesions using the support vector machine. Although they used a supervised ML method, they also aimed to obtain the cancer lesion 'region'. Providing a 'line' rather than a 'region' using a CAD system is meaningful in enriching the knowledge of both inexperienced and experienced endoscopists.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In determining the horizontal extent of early gastric mucosal cancer using M-NBI, it is necessary to identify the DL to differentiate the mucosal patterns between cancer lesions and background mucosa. Consequently, the DL exhibits a complex geometric shape, as illustrated in **Figure 9**. CASE 4 in **Figure 9(a)** included depressed lesions, and the endoscopist indicated the DL on the boundary between the depressed lesion and background mucosa. Moreover, **Figure 9(c)** indicates that the DL was drawn by tracing the light blue crest in CASE 3. In both cases, the endoscopists drew the DLs to trace the structural change in the white zone, and the DLs resulted in an intricate shape being formed. Notably, our proposed system could provide these irregular and complex shapes precisely, as illustrated in **Figures 9(b)** and **(d)**.

However, this study exhibits certain limitations. First, the small sample size (n = 4) is problematic. Our system needs to be tested with a larger sample size to improve its reliability and robustness. This work focused on the proposal and development of a general framework for an automatic DL determination algorithm based on a data clustering method. Second, each DL may contain an individual bias because a single endoscopist drew it for each image. Although the DLs were assessed by multiple doctors in the present data, the other DL candidates should be provided by various endoscopists to reduce such bias. Finally, pathological evaluation has not been carried out in this study. Biopsies and the use of their results as ground truth for evaluating our proposed system are necessary to improve the system reliability and to proceed to clinical trials. Future work will focus on these issues. However, irrespective of these limitations, the current results demonstrate the effectiveness of our proposed method, and will contribute the development of gastroenterology.

5. CONCLUSIONS

We developed a new CAD system that can automatically determines the DL on a given M-NBI image through an ML method. In the proposed system, the boundaries of the clusters within M-NBI images are extracted as the DL candidates. The performance evaluation of the system on actual M-NBI images

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

revealed that it can represent the complex shapes of the DLs, similarly to the DLs identified by experienced doctors. The identification of the DL is often quite a difficult task for inexperienced endoscopists. Therefore, our proposed system, which can provide a variety of DL candidates, may support the training of inexperienced doctors, as well as enriching the knowledge of experienced doctors.

FIGURE LEGENDS

Figure 1. Framework of proposed method.

Figure 2. Examples of cluster boundaries identified after global clustering: (a) superpixels identified by local clustering, (b) clusters of superpixels (k = 2), where each cluster is painted with a different color, (c) cluster boundaries (k = 2), (d) clusters of superpixels (k = 3), (e) cluster boundaries (k = 3), (f) clusters of superpixels (k = 4), and (g) cluster boundaries (k = 4).

Figure 3. Image with DL drawn by experienced endoscopist: (a) raw M-NBI image and (b) M-NBI image with DL.

Figure 4. Schematic of performance metric calculation. The blue and green lines indicate the DL candidates provided by the endoscopists and proposed method, respectively. The black dashed arrows illustrate the Euclidean distance calculation from each pixel of the DL provided by the endoscopists to the nearest pixel of the DL candidates provided by the proposed method.

Figure 5. DL candidates provided by proposed method at different cluster number settings (k = 2, 3, 4, 5, and 6) for 16 M-NBI images obtained from 4 cases. The leftmost images present the DLs provided by the endoscopists. The Euclidean distance between the true DL and DL candidates provided by the proposed method is indicated below each image at different k values. The average, minimum, and maximum distance were calculated with different k settings for each case.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Figure 6. Dependence of Euclidean distance between DLs provided by endoscopists and proposed method on number of clusters. Each black dashed line indicates the Euclidean distance for each case at each *k* value. The average distance for all of the images at each *k* value was calculated and is plotted with a blue point, and the points are connected with a dashed blue line.

Figure 7. DL candidates identified by proposed method for CASE 4. The upper images present the DL candidates at each *k* value. The yellow and green lines indicate the DLs provided by the endoscopists and proposed system, respectively. The lower plot shows the dependence of the Euclidean distance between the DLs provided by the endoscopists and proposed method on different *k* values.

Figure 8. DL candidates identified by proposed method for CASE 3. The upper images present the DL candidates at each *k* value. The lower plot indicates the dependence of the Euclidean distance between the DLs provided by the endoscopist and proposed method on different *k* values.

Figure 9. Comparison of DLs provided by endoscopist and proposed method. The top two images (a) and (b) indicate the DLs identified by the endoscopist and proposed system, respectively, for CASE 4, whereas the bottom two images (c) and (d) are for CASE 3.

REFERENCES

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[1] The Global Cancer Observatory - Globocan 2018, The International Agency for Research on Cancer (IARC), viewed 11 August 2020, <https://gco.iarc.fr/today/data/factsheets/cancers/7-Stomach-fact-sheet.pdf>.

[2] T. Sano, D.G. Coit, H.H. Kim, F. Roviello, P. Kassab, C. Wittekind, Y. Yamamoto and Y. Ohashi, "Proposal of a new stage grouping of gastric cancer for tnm classification: International gastric cancer association staging project," Gastric cancer, vol. 20, no. 2, pp. 217–225, 2017.

[3] H. Katai, T. Ishikawa, K. Akazawa, Y. Isobe, I. Miyashiro, I. Oda, S. Tsujitani, H. Ono, S. Tanabe, T. Fukagawa, et al., "Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001–2007)," Gastric Cancer, vol. 21, no. 1, pp. 144–154, 2018.

[4] K. Yao, "Clinical application of magnifying endoscopy with narrow-band imaging in the stomach," Clinical endoscopy, vol. 48, no. 6, p.481, 2015.

[5] Y. Ezoe, M. Muto, N. Uedo, H. Doyama, K. Yao, I. Oda, K. Kaneko, Y. Kawahara, C. Yokoi, Y. Sugiura, et al., "Magnifying narrowband imaging is more accurate than con- ventional white-light imaging in diagnosis of gastric mucosal cancer," Gastroenterology, vol. 141, no. 6, pp. 2017–2025, 2011.

[6] K. Yao, G. Anagnostopoulos and K. Ragunath, "Magnifying endoscopy for diagnosing and delineating early gastric cancer," Endoscopy, vol. 41, no. 05, pp. 462–467, 2009.

[7] T. Gotoda, H. Yamamoto and R.M. Soetikno, "Endoscopic submucosal dissection of early gastric cancer," Journal of gastroenterology, vol. 41, no. 10, pp. 929–942, 2006.

[8] M. Tanaka, H. Ono, N. Hasuike and K. Takizawa, "Endoscopic submucosal dissection of early gastric cancer," Digestion, vol. 77, no. Suppl. 1, pp. 23–28, 2008.

[9] N. Kakushima, H. Ono, M. Tanaka, K. Takizawa, Y. Yamaguchi and H. Matsubayashi, "Factors related to lateral margin positivity for cancer in gastric specimens of endoscopic submucosal dissection," Digestive Endoscopy, vol. 23, no. 3, pp. 227–232, 2011.

[10] M. Sekiguchi, H. Suzuki, I. Oda, S. Abe, S. Nonaka, S. Yoshinaga, H. Taniguchi, S. Sekine, R. Kushima and Y. Saito, "Risk of recurrent gastric cancer after endoscopic resection with a positive lateral margin," Endoscopy, vol. 46, no. 04, pp. 273–278, 2014.

[11] T.K. Kim, G.H. Kim, B.E. Lee, T.Y. Jeon, D.H. Kim, H.J. Jo, G. Am Song, et al., "Risk factors for local recurrence in patients with positive lateral resection margins after endoscopic submucosal dissection for early gastric cancer," Surgical endoscopy, vol. 29, no. 10, pp. 2891–2898, 2015.

[12] A. Imagawa, H. Okada, Y. Kawahara, R. Takenaka, J. Kato, H. Kawamoto, S. Fujiki, R. Takata, T. Yoshino and Y. Shiratori, "Endoscopic submucosal dissection for early gastric cancer: results and degrees of technical difficulty as well as success," Endoscopy, vol. 38, no. 10, pp. 987–990, 2006.

medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

[13] T. Sugimoto, M. Okamoto, Y. Mitsuno, S. Kondo, K. Ogura, T. Ohmae, H. Mizuno, S. Yoshida, Y. Isomura, Y. Yamaji, et al., "Endoscopic submucosal dissection is an effective and safe therapy for early gastric neoplasms: a multicenter feasible study," Journal of clinical gastroenterology, vol. 46, no. 2, pp. 124–129, 2012.

[14] H. Iizuka, S. Kakizaki, N. Sohara, Y. Onozato, H. Ishihara, S. Okamura, H. Itoh and M. Mori, "Stricture after endoscopic submucosal dissection for early gastric cancers and adenomas," Digestive Endoscopy, vol. 22, no. 4, pp. 282–288, 2010.

[15] Y. Horie, T. Yoshio, K. Aoyama, S. Yoshimizu, Y. Horiuchi, A. Ishiyama, T. Hirasawa, T. Tsuchida, T. Ozawa, S. Ishihara, et al., "Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks," Gastrointestinal endoscopy, vol. 89, no. 1, pp. 25–32, 2019.

[16] T. Hirasawa, K. Aoyama, T. Tanimoto, S. Ishihara, S. Shichijo, T. Ozawa, T. Ohnishi, M. Fujishiro, K. Matsuo, J. Fujisaki, et al., "Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images," Gastric Cancer, vol. 21, no. 4, pp. 653–660, 2018.

[17] T. Yasuda, T. Hiroyasu, S. Hiwa, Y. Okada, S. Hayashi, Y. Nakahata, Y. Yasuda, T. Omatsu, A. Obora, T. Kojima, et al., "Potential of automatic diagnosis system with linked color imaging for diagnosis of helicobacter pylori infection," Digestive Endoscopy, pp.••–••, 2019.

[18] M. Misawa, S.-e. Kudo, Y. Mori, H. Nakamura, S. Kataoka, Y. Maeda, T. Kudo, T. Hayashi, K. Wakamura, H. Miyachi, et al., "Characterization of colorectal lesions using a computer-aided diagnostic system for narrow-band imaging endocytoscopy," Gastroenterology, vol. 150, no. 7, pp. 1531–1532, 2016.

[19] S. Shichijo, S. Nomura, K. Aoyama, Y. Nishikawa, M. Miura, T. Shinagawa, H. Takiyama, T. Tanimoto, S. Ishihara, K. Matsuo, et al., "Application of convolutional neural networks in the diagnosis of helicobacter pylori infection based on endoscopic images," EBioMedicine, vol. 25, pp. 106–111, 2017.

[20] Y. Mori, S.-e. Kudo, H.E. Mohmed, M. Misawa, N. Ogata, H. Itoh, M. Oda and K. Mori, "Artificial intelligence and upper gastrointestinal endoscopy: current status and future perspective," Digestive Endoscopy, vol. 31, no. 4, pp. 378–388, 2019.

[21] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., & Süsstrunk, S. (2012). SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, *34*(11), 2274-2282.

[22] G. Wyszecki and W.S. Stiles, Color science, Vol.8, Wiley New York, 1982.

[23] M.K. Agoston and M.K. Agoston, Computer graphics and geometric modeling, Vol.1, Springer, 2005.

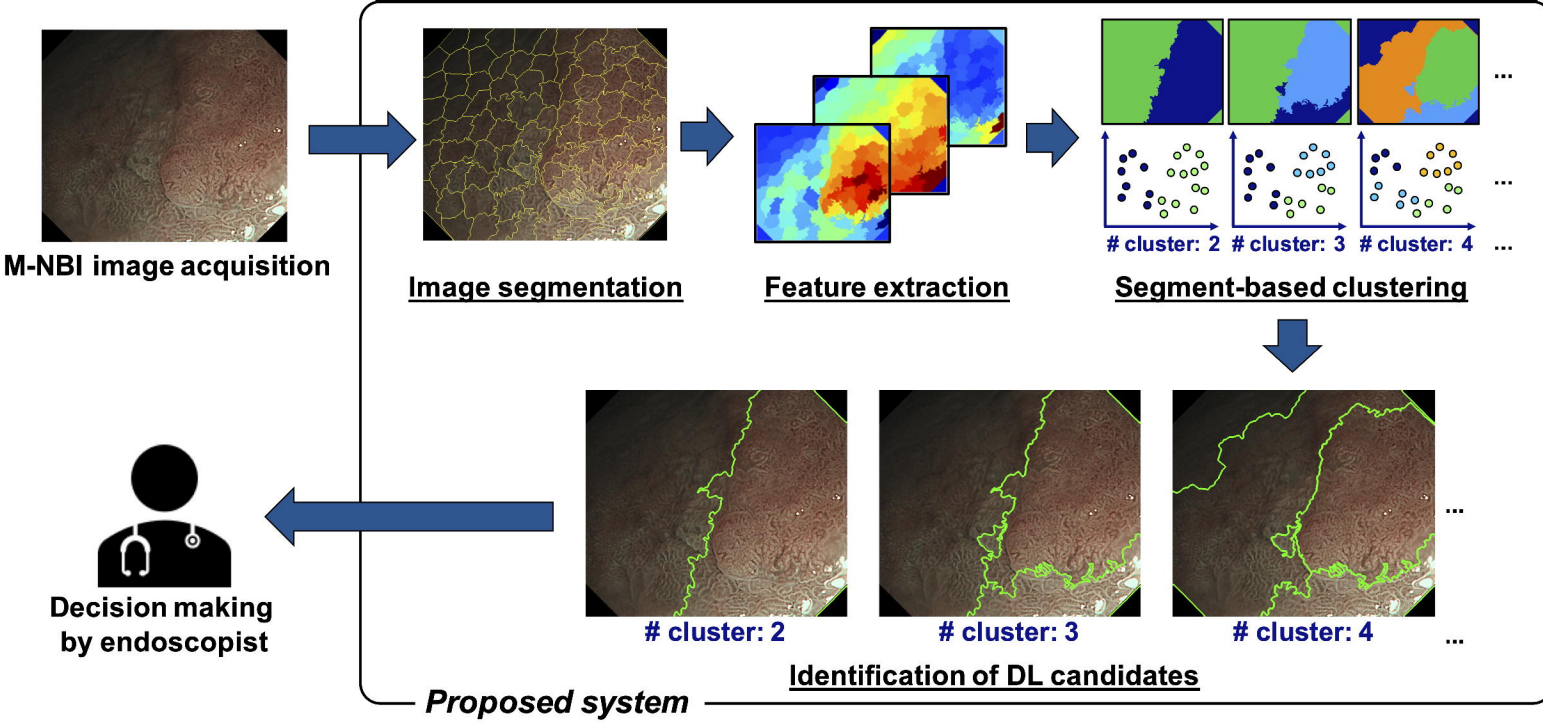
[24] Yao, K., Anagnostopoulos, G. K., & Ragunath, K. (2009). Magnifying endoscopy for diagnosing and delineating early gastric cancer. *Endoscopy*, *41*(05), 462-467.

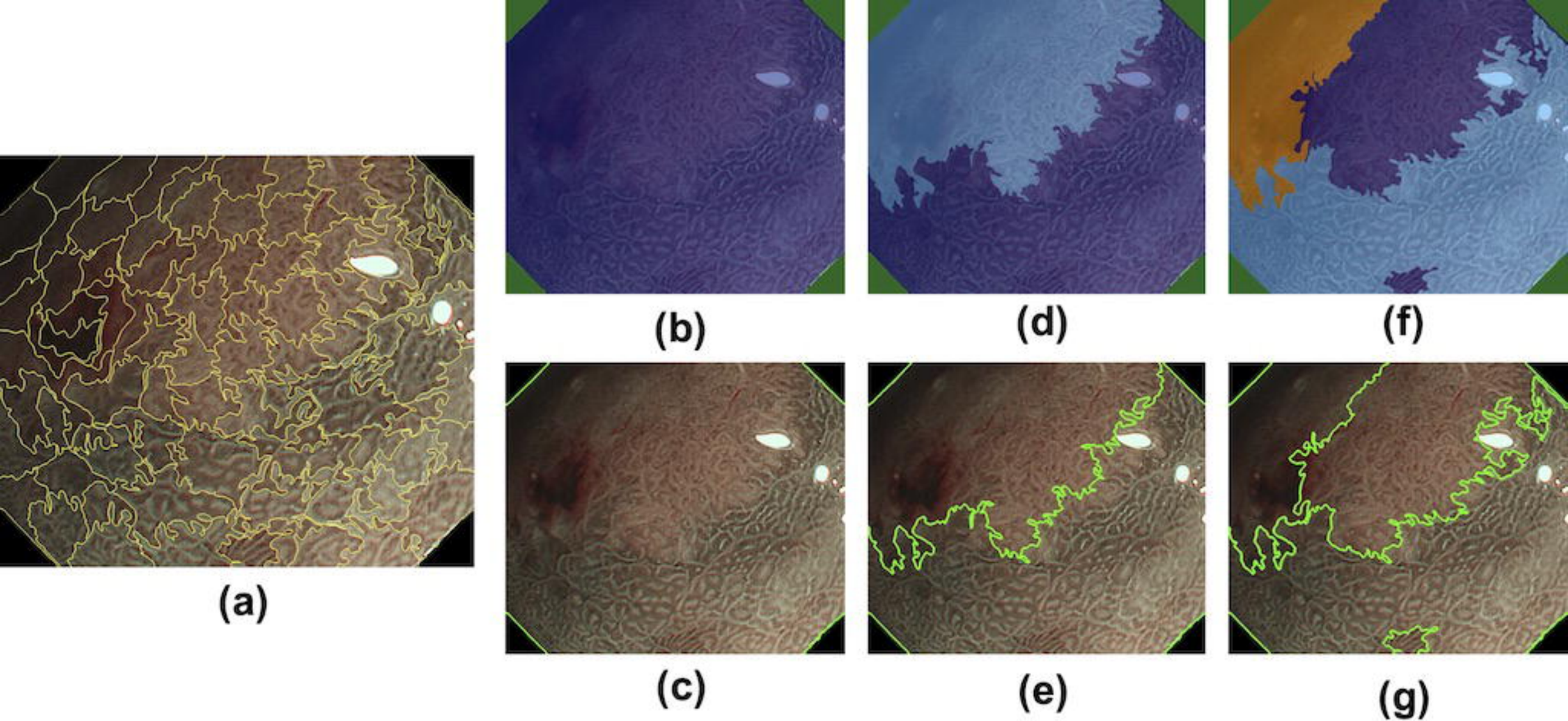
medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

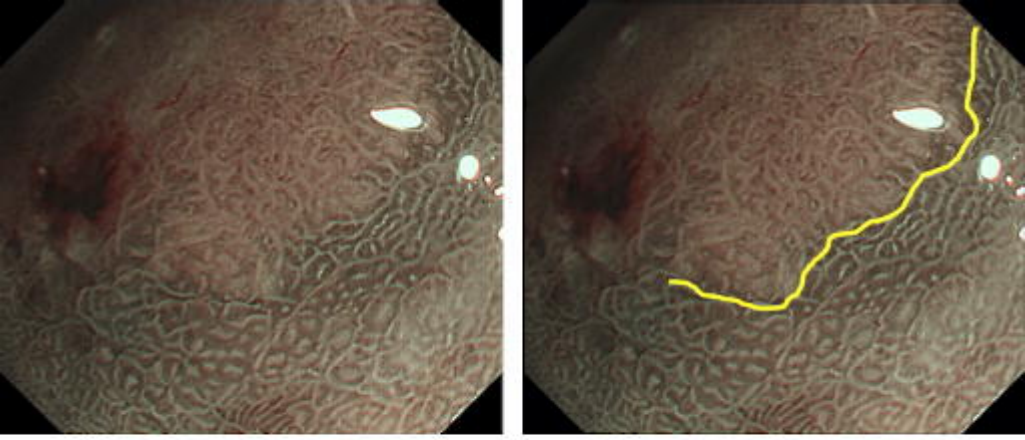
[25] Wu, Y., Zhou, Y., Saveriades, G., Agaian, S., Noonan, J. P., & Natarajan, P. (2013). Local Shannon entropy measure with statistical tests for image randomness. *Information Sciences*, *222*, 323-342.

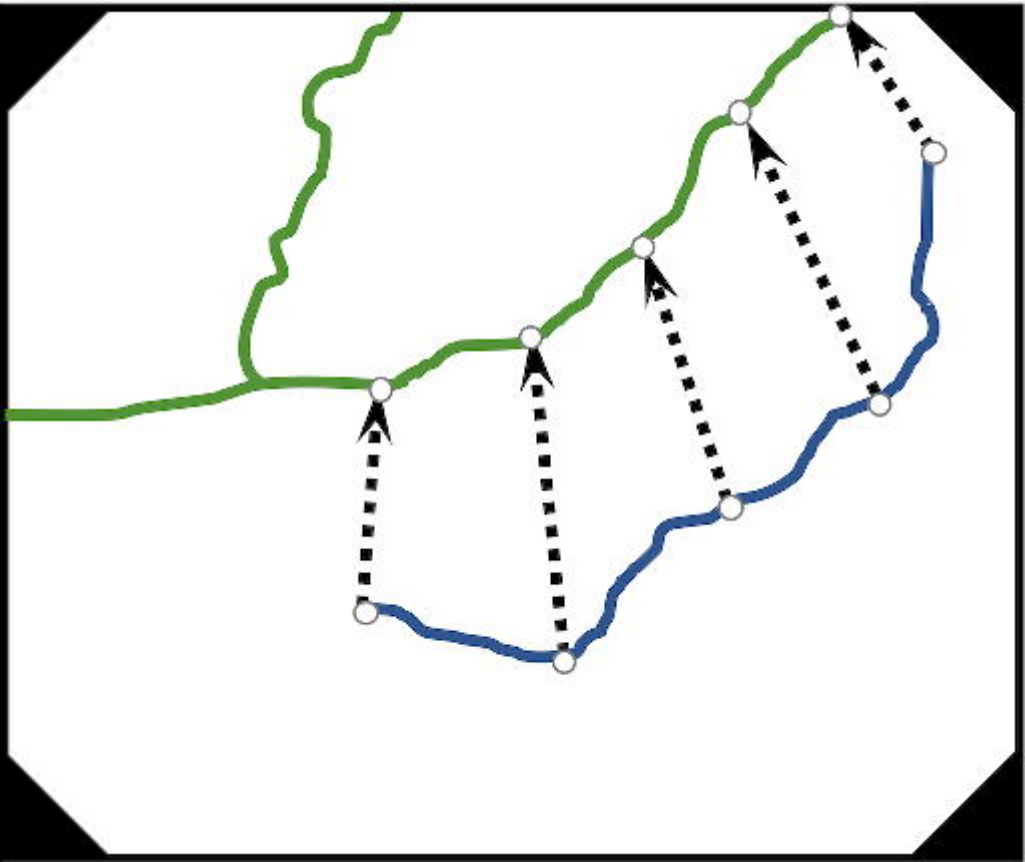
[26] L. Li, Y. Chen, Z. Shen, X. Zhang, J. Sang, Y. Ding, X. Yang, J. Li, M. Chen, C. Jin, et al., "Convolutional neural network for the diagnosis of early gastric cancer based on magnifying narrow band imaging," Gastric Cancer, vol. 23, no. 1, pp. 126–132, 2020.

[27] T. Kanesaka, T.-C. Lee, N. Uedo, K.-P. Lin, H.-Z. Chen, J.-Y. Lee, H.-P. Wang and H.-T. Chang, "Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging," Gastrointestinal endoscopy, vol. 87, no. 5, pp. 1339–1344, 2018.

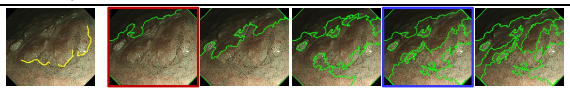
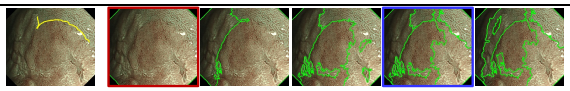
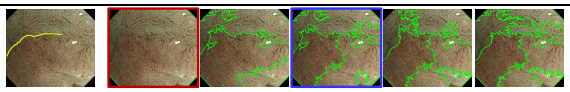
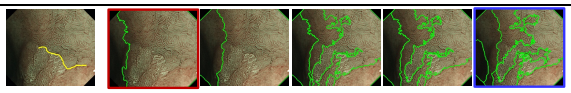
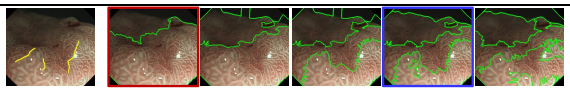
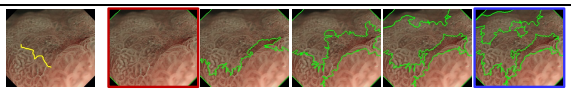
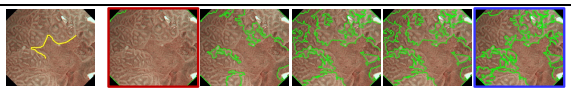
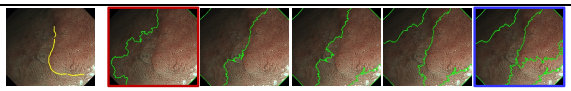
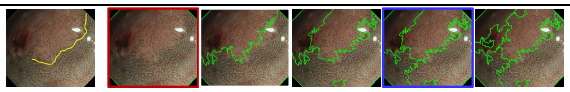
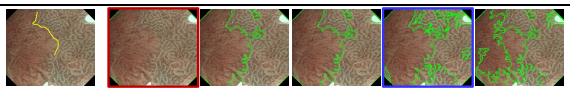
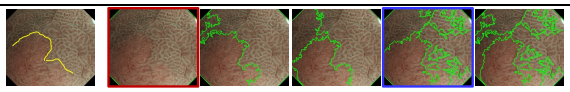
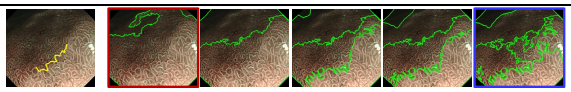
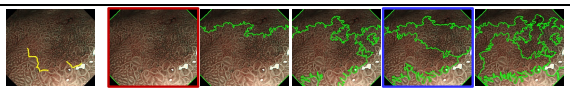
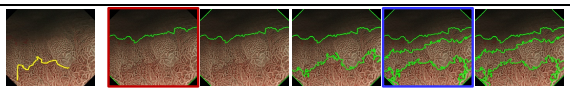
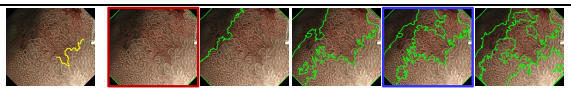
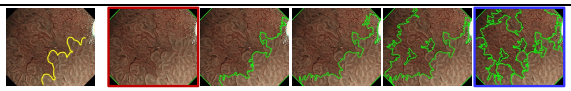

M-NBI image acquisition
Image segmentation
Feature extraction
cluster: 2
cluster: 3
cluster: 4
Segment-based clustering
cluster: 2
cluster: 3
cluster: 4
Identification of DL candidates
Decision making by endoscopist
Proposed system

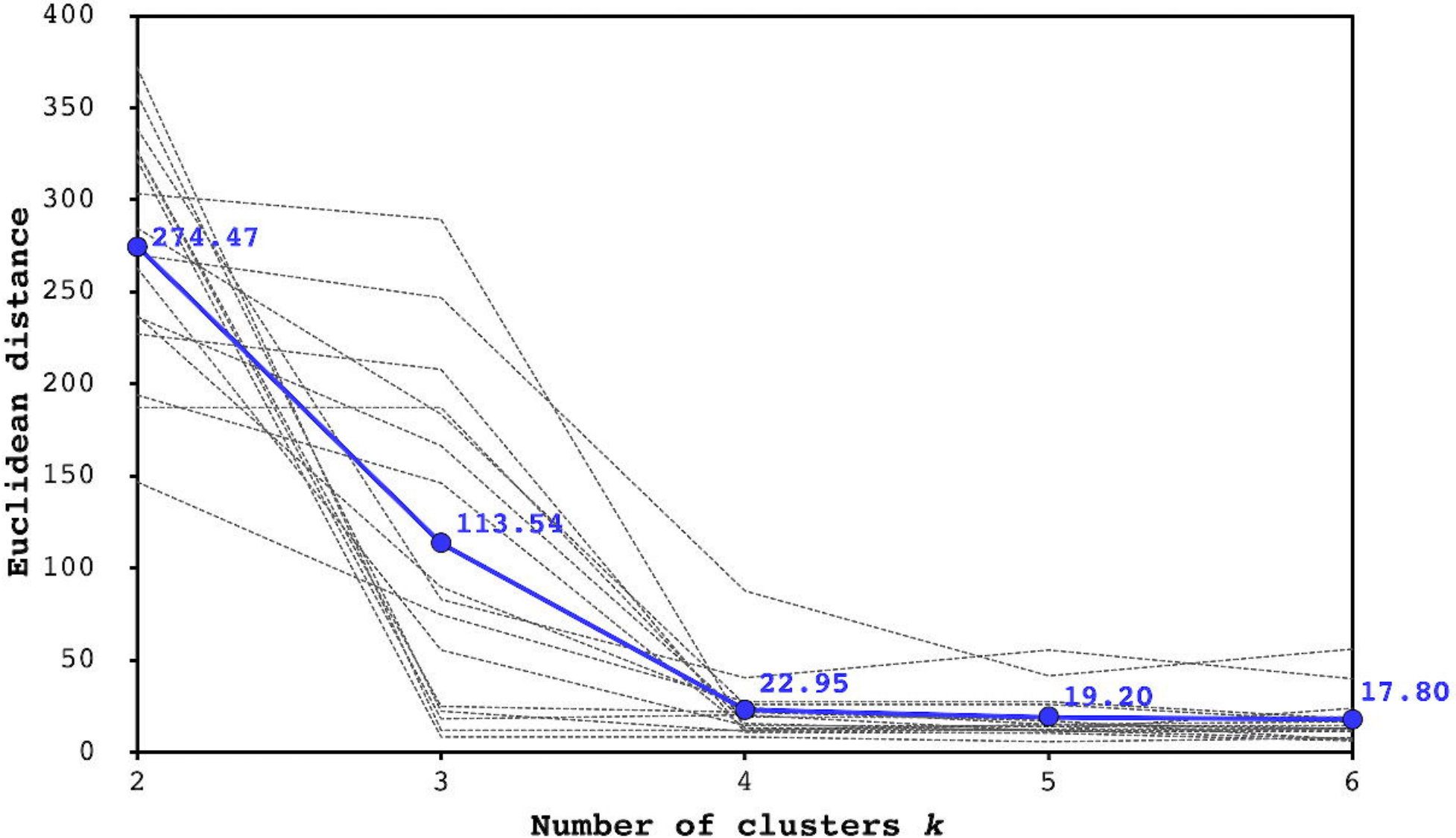

(a)
(b)
(c)
(d)
(e)
(f)
(g)

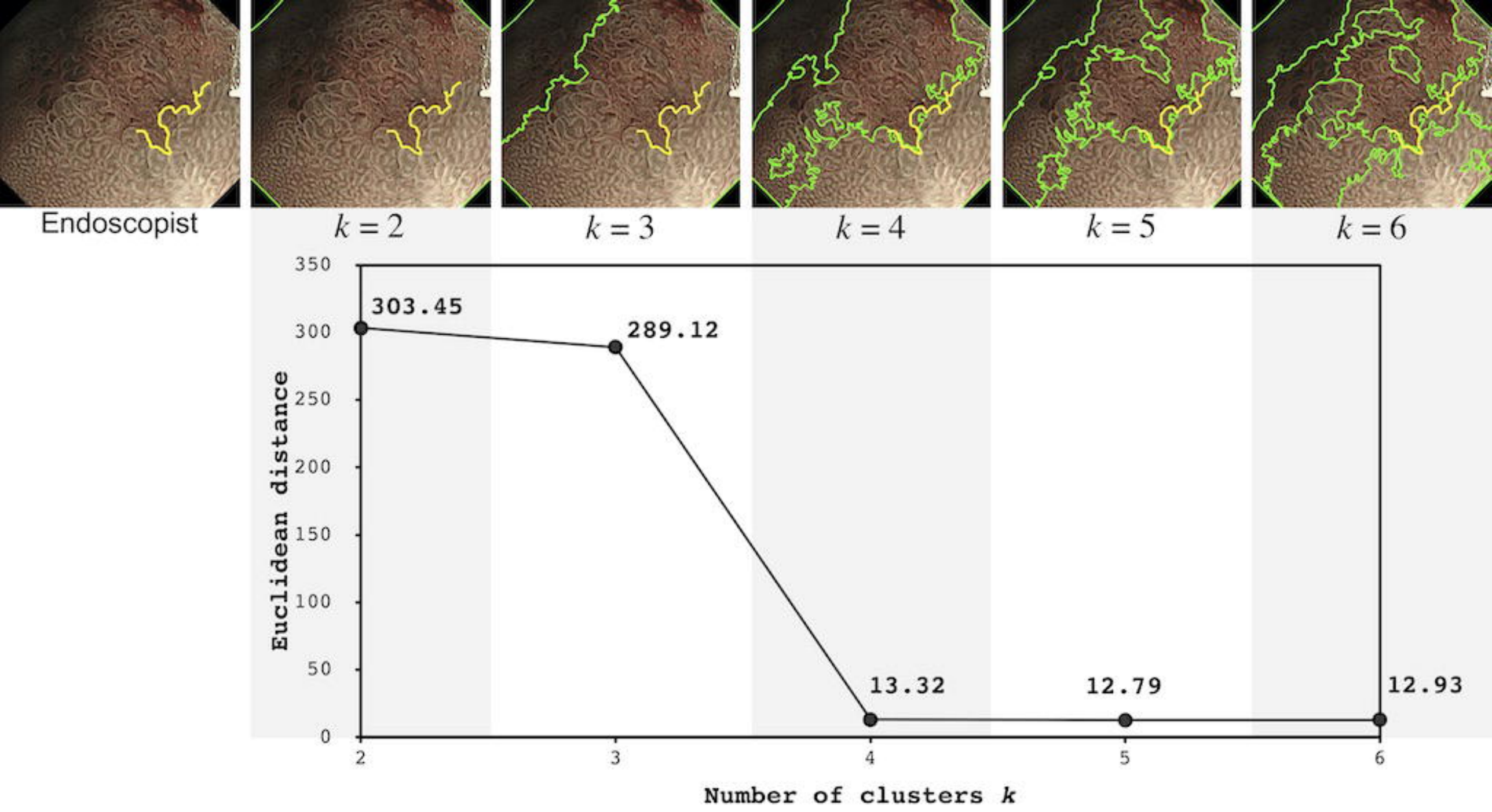

(a)
(b)

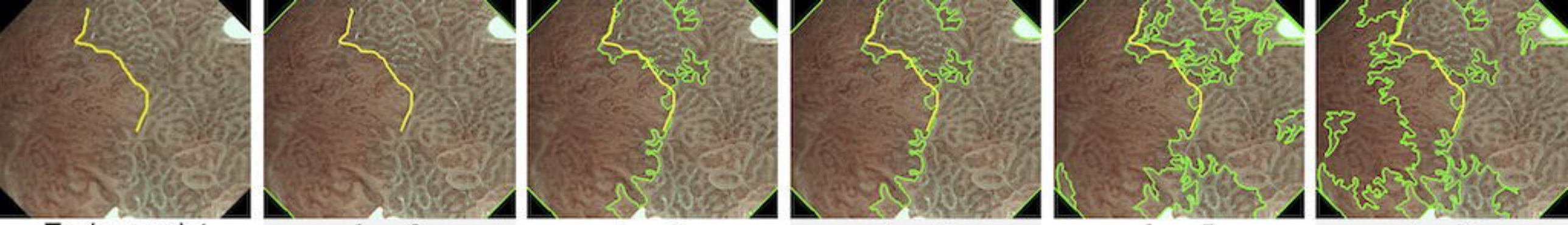


medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

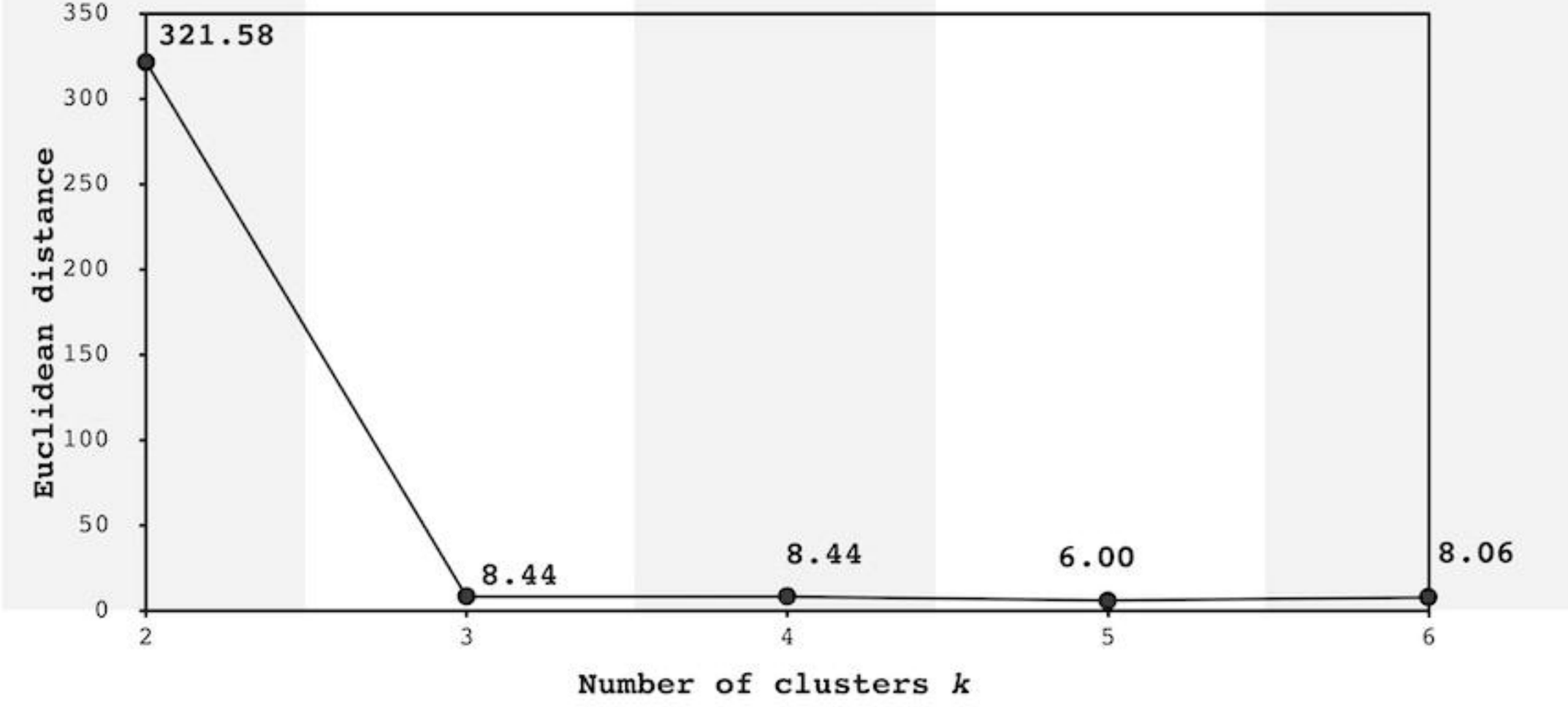
Case		Endoscopist	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	Average	Min.	Max.
CASE 1	1		**236.33**	166.34	19.16	**18.94**	19.31	92.02	18.94	236.33
	2		**236.91**	89.64	24.25	**14.96**	16.80	76.51	14.96	236.91
	3		**371.87**	11.94	**11.90**	15.07	12.92	84.74	11.90	371.87
	4		**226.94**	207.91	10.87	10.48	**6.47**	92.53	6.47	226.94
		Case-wise average						86.45	**13.07**	268.01
CASE 2	1		**193.83**	146.33	12.98	**12.01**	23.79	77.79	12.01	193.83
	2		**338.44**	83.16	40.75	55.50	**40.21**	111.61	40.21	338.44
	3		**357.47**	22.28	11.76	11.87	**7.64**	82.20	7.64	357.47
		Case-wise average						90.53	**19.95**	296.58
CASE 3	1		**147.03**	74.77	27.45	27.45	**18.45**	59.03	18.45	147.03
	2		**263.01**	55.80	14.69	**14.29**	14.42	72.44	14.29	263.01
	3		**321.58**	8.44	8.44	**6.00**	8.06	70.50	6.00	321.58
	4		**326.10**	18.13	20.12	**11.36**	12.43	77.63	11.36	326.10
		Case-wise average						69.90	**12.53**	264.43
CASE 4	1		**284.42**	183.74	26.05	25.79	**16.77**	107.35	16.77	284.42
	2		**270.03**	246.91	87.75	**41.65**	56.08	140.48	41.65	270.03
	3		**187.42**	187.42	15.81	**11.59**	11.78	82.80	11.59	187.42
	4		**303.45**	289.12	13.32	**12.79**	12.93	126.32	12.79	303.45
	5		**326.76**	24.76	21.97	17.39	**6.71**	79.52	6.71	326.76
		Case-wise average						107.29	**17.90**	274.42


400
350
300
250
200
150
100
50
0
Euclidean distance
274.47
113.54
22.95
19.20
17.80
2
3
4
5
6
Number of clusters k

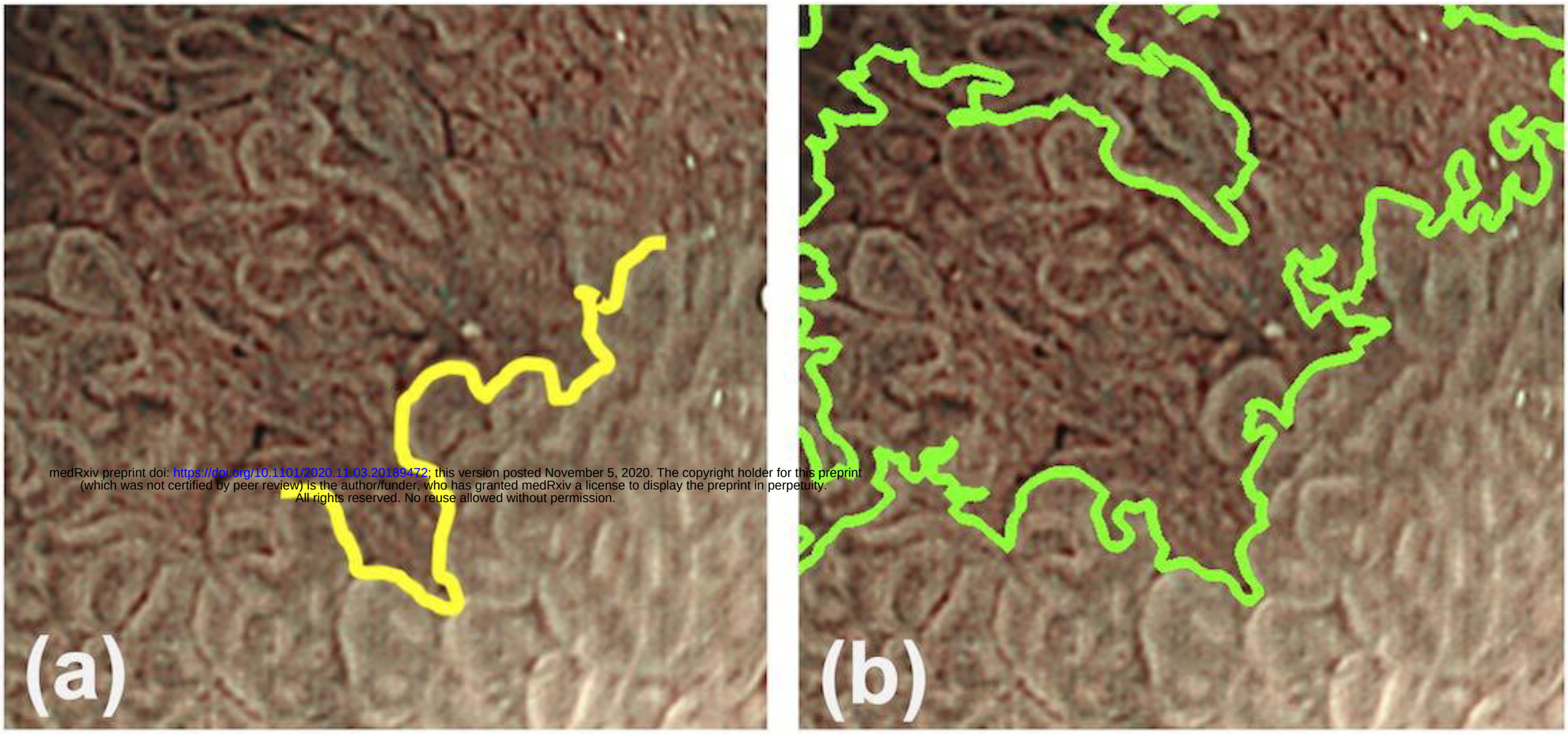

Endoscopist
k = 2
k = 3
k = 4
k = 5
k = 6
Euclidean distance
Number of clusters k
303.45
289.12
13.32
12.79
12.93
0
50
100
150
200
250
300
350
2
3
4
5
6



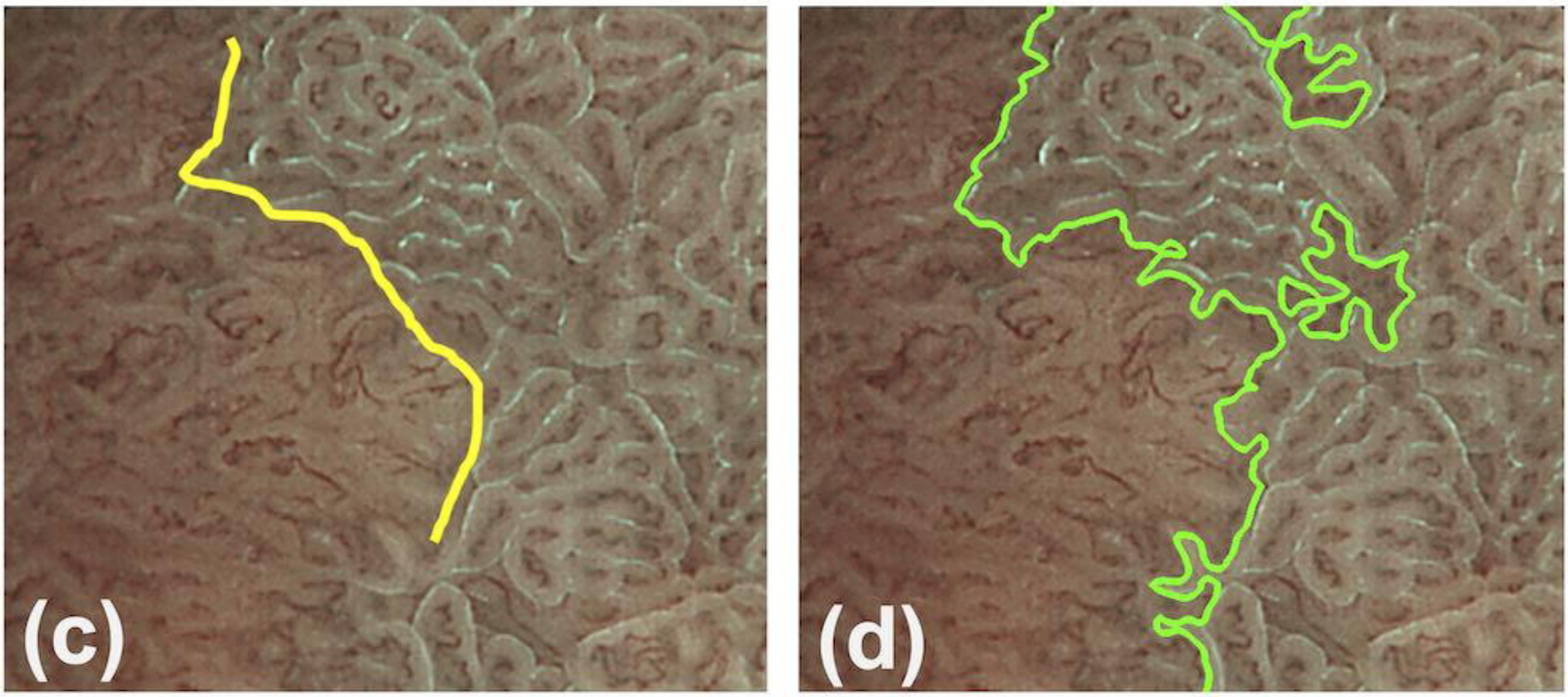
Endoscopist $k = 2$ $k = 3$ $k = 4$ $k = 5$ $k = 6$



medRxiv preprint doi: https://doi.org/10.1101/2020.11.03.20189472; this version posted November 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.



CASE 4 (left: endoscopist, right: proposed system)



CASE 3 (left: endoscopist, right: proposed system)